Notes on Interference and Diffraction

September 13, 2001

Cornell University

Department of Physics

Physics 214 September 13, 2001

The purpose of these notes is to understand interference in $d > 1$ dimensions, particularly diffraction and diffraction gratings as appear in Lab 3. The most important parts to go over before doing the lab are the interpretation sections, Sections 3.2.2, 3.3.2, 4.2 & 5.2, and the formulas which go along with them, Eqs. 13, 17, 19 & 20.

Contents

List of Figures

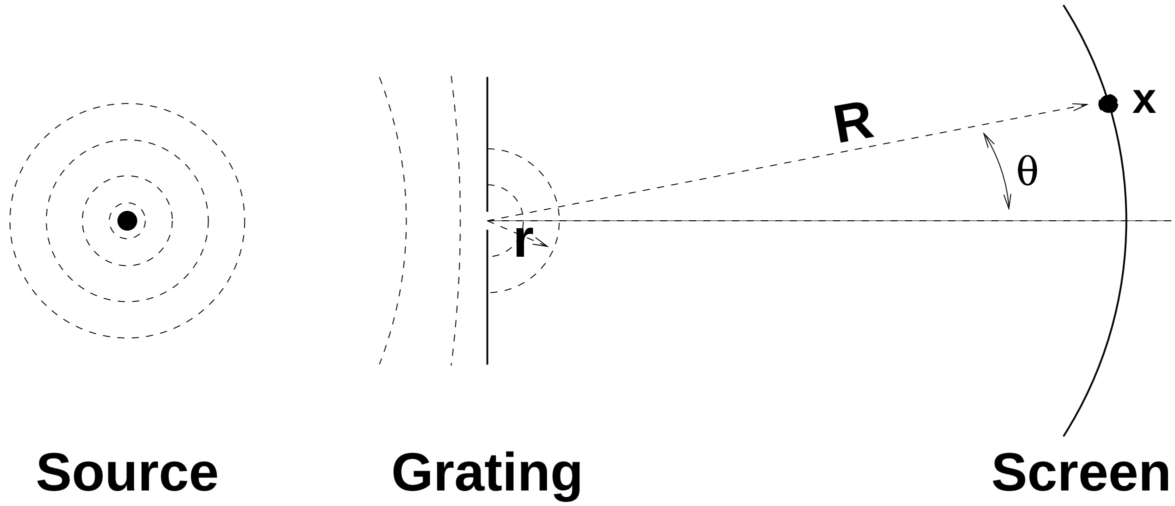

Figure 1: Single-slit experiment

1 Overview

Section 2 begins by considering the simplest case of a diffraction grating, a single, infinitesimally narrow slit. The purpose of this section is to introduce the primary concepts and approximations used in the later sections. If you are happy accepting Eq. 8 on faith, you can skip Section 2 and begin directly with Section 3.2.

Section 3 studies the phenomenon of *interference*, which happens in cases where there are more than one slit. The central result of these notes is the N-slit diffraction pattern (Eq. 17) from this section. From this equation, one may understand the two slit pattern ($N = 2$) as a special case, or deal with any larger number of slits. In fact, from this one result, one may also understand the pattern for a single slit of finite width (Section 4) and for the more complex case of N slits of *finite* width (Section 5).

2 Pattern from a single slit of infinitesimal width

We begin by considering the wave pattern created on the observation screen in the experiment which Figure 1 shows. There is a source of waves far to the left. This waves then pass through a *diffraction grating*, which blocks most of the wave and letting only a small part through, in this case that which passes through a single, very narrow slit. What passes through the grating is then collected on the observation screen. Our goal is to understand what is observed on this screen.

This section explains the reasoning behind the basic results upon which our analyses rest: (a) that the intensity from a single slit depends only on distance from the slit, (b) the amplitude of the waves from a single slit is *inversely proportional* to distance from the slit, (c) the phase of the waves from a single slit (in the complex representation) goes like e^{ikr} where r is distance from the slit.

2.1 Meaning of infinitesimal

In physics (or politics, for that matter), one should always be suspicious when told that something is "small" or "large" and ask, "Small or large *compared to what*!?!" In this case, we are speaking relative to the wavelength λ of the waves going through the slit. As long as the slit is very narrow *compared to* λ, then our results below will be valid.

2.2 Relation between intensity and amplitude

Now, the *intensity* I ("brightness" in the case of light) of the pattern at the observation screen is a just a measure of the amount of energy arriving at each point $\vec{x}$ on the screen per unit time. In a real experiment, the screen is a two-dimensional surface. Thus, the precise measure we use for the intensity is the average energy arriving per unit time *per unit area*, or, simply, average power per unit area. To relate this to the solution for the waves $y(\vec{x}, t)$, we can generalize from the result we know for power on a string,

$$P(x) = -\tau \frac{\partial y}{\partial x} \frac{\partial y}{\partial t}. \tag{1}$$

It turns out that all phenomena (such as sound or light) which obey the wave-equation obey a very similar formula. The primary difference in going from waves on a string to any other type of wave is that the constant τ (which measures the tension of the string) is replaced by some other constant characterizing the medium of propagation for the waves. If we consider traveling sinusoidal waves, then the solution will have the general form $y(x,t) = A\cos(kx - \omega t)$. Then, using Eq. 1 we will find a result of the following form for the power. (Here, $\langle q \rangle$ indicates the average value of the quantity q.)

$$\begin{aligned} I &= \langle -\tau \left(-kA\sin(kx - \omega t)\right) \left(\omega A \sin(kx - \omega t)\right) \rangle \\ &= \tau k \omega A^2 \left\langle \sin^2(kx - \omega t) \right\rangle. \end{aligned}$$

Note that, at any instant in time t, the power may be different: it varies with the oscillations of $\sin^2(kx - \omega t)$. This is why we define intensity as the *time average* of the arriving power. Finally, because the time average of $\sin^2$ or $\cos^2$ is just 1/2, we have

$$\begin{aligned} I &= \frac{1}{2} \tau k \omega A^2 \\ &\equiv \alpha A^2. \end{aligned} \tag{2}$$

Note that in the last line we replaced the factor $(1/2)\tau\omega k$ with a general proportionality constant α. We do this because for waves which are not waves on a string, we end up with factors other than τ to describe the medium of propagation. What is the same for all types of waves, the *main point* here, is that the *intensity* is proportional to the *square of the amplitude*.

2.3 Form of solution from a narrow slit

Because we consider a slit which is much narrower than the wavelength of the waves passing through it, *physically* there is no way for the waves to distinguish the slit from a tiny point source of waves. This is known as *Huygen's principle*. Mathematically, it is the same principle which makes the electric potential at a distance r far away from a distribution of charges which are spread out over a small distance $d << r$ go just like the field from a single point charge, kQ/r, regardless of exactly how the charges are arranged.

2.3.1 Amplitude

If the slit in Figure 1 were truly a point source of waves, then we expect the waves to come out equally in all directions (as indicated in the figure). Moreover, the average total power crossing any half-sphere of radius r (and thus area $a = (1/2)4\pi r^2 = 2\pi r^2$) must equal the total power P put out by the source. Because the intensity gives the average power per unit time per unit area, we must have, using Eq.2 for the intensity,

$$\begin{aligned} P &= (I)(a) \\ &= (\alpha A^2)(2\pi r^2) \\ &\Rightarrow \\ A &= \frac{\sqrt{P/(2\pi\alpha)}}{r}, \end{aligned} \tag{3}$$

where α is the proportionality constant from Eq. 2. The basic fact that we have learned the *amplitude* of the waves from the slit to be *inversely* proportional to the distance from the slit.

2.3.2 Phase

To complete the specification of the solution at any point $\vec{x}$, we need not only the amplitude, but also the phase. For waves traveling out in the radial direction away from the point source, we expect a form just like the string, $y(r,t) = A\cos(kr - \omega t + \phi)$, where ϕ is some phase. Given what we know of the amplitude from Eq. 3, the solution at any point $\vec{x}$ at distance $R = |\vec{x}|$ from the slit is then

$$\begin{aligned} y(\vec{x}, y) &= \frac{\sqrt{P/(2\pi\alpha)}}{R}\cos(kR - \omega t + \phi) \qquad (4) \\ &= \text{Re}\left\{\frac{\sqrt{P/(2\pi\alpha)}}{R}e^{i\phi}e^{ikR}e^{-i\omega t}\right\} \end{aligned}$$

Note that, we here have written the answer also in complex form, $y(\vec{x}, t) = \text{Re}\left\{A(\vec{x})e^{-i\omega t}\right\}$ which describes simple harmonic oscillations of complex amplitude

$$A(\vec{x}) \equiv \frac{\sqrt{P/(2\pi\alpha)}}{R}e^{i\phi}e^{ikR}. \qquad (5)$$

This will be useful in future sections.

2.4 Intensity pattern from a single slit

Armed now with the solution from a single slit (Eq. 4) and the relation between amplitude and intensity (Eq. 3), we are ready to compute the intensity pattern from a single slit. For simplicity we take the observation screen to be curved like the inside of a large dome of constant radius R (as in Figure 1). As long as the radius of curvature of the dome R is much larger than the actual size of the screen, this will have no meaningful effect on our results.

For the intensity, we have now

$$\begin{aligned} I &= \alpha A^2 \\ &= \alpha\left|A(\vec{x})\right|^2 \\ &= \alpha\left|\frac{\sqrt{P/(2\pi\alpha)}}{R}e^{i\phi}e^{ikR}\right|^2 \\ &= \alpha\left|\frac{\sqrt{P/(2\pi\alpha)}}{R}\right|^2\left|e^{i\phi}\right|^2\left|e^{ikR}\right|^2 \\ &= \frac{P}{2\pi R^2}, \qquad (6) \end{aligned}$$

where we have substituted Eq. 5 for $A(\vec{x})$ and used the mathematical facts that $|\underline{\text{ab}}\ldots\underline{\text{c}}|^2 = |\underline{\text{a}}|^2|\underline{\text{b}}|^2\ldots|\underline{\text{c}}|^2$ for any product of complex numbers and that $|e^{ikR}| = \sqrt{\cos^2(kR) + \sin^2(kR)} = 1$.

The main lesson we learn from Eq. 6 is that for a single slit, the pattern on an observation screen at distance R from the slit is constant and uniform, independent of the angle of observation θ, as Figure 2 illustrates.

3 Multiple slits

We now consider what happens in the presence of more than one slit.

3.1 Shortcut

This section focuses on a very useful algebraic simplification we can make to analyze interference patterns. The main lesson is that the intensity at point $\vec{x}$ is just $I = |\sqrt{I_1}e^{i(kr_1+\phi_1)} + \ldots + \sqrt{I_N}e^{i(kr_N+\phi_N)}|^2$, where I_n

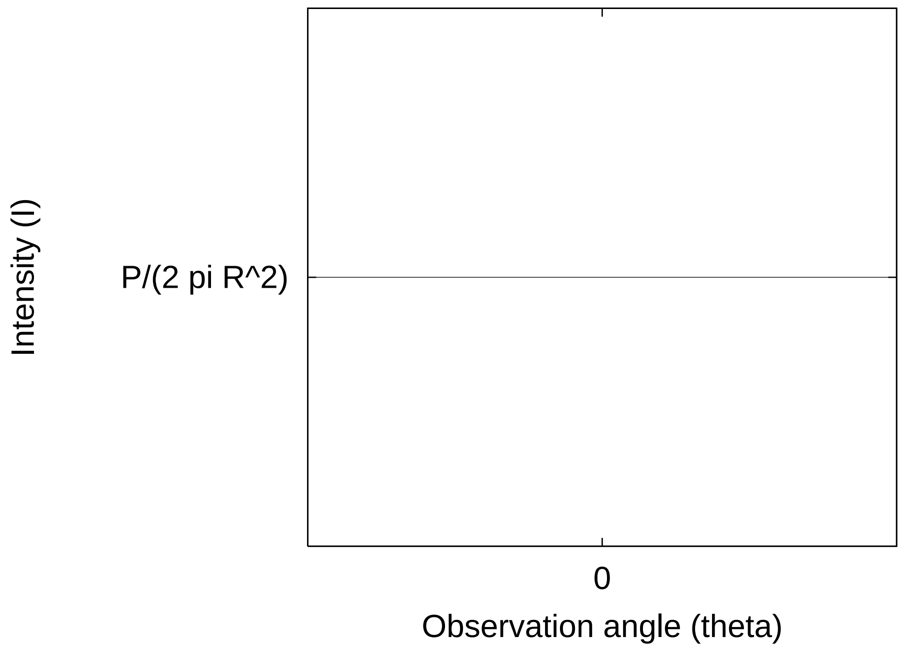

Figure 2: Intensity as a function of observation angle for a single-slit is constant

is the intensity that slit n would give if it were the only slit, r_n is the distance from slit n to the observation point, and ϕ_n is the phase of the waves as they just emerge from slit n. In fact, if you know just this result, you can derive the intensity for any pattern of slits!

To understand this shortcut, we begin with Eqs. 4 & 5, which tell us that the solution coming from the n^{th} slit, S_n, arriving at point $\vec{x}$ on the screen is

$$y_n(\vec{x}, t) = \text{Re}\left\{A_n(\vec{x})e^{-i\omega t}\right\},$$

with

$$A_n(\vec{x}) \equiv \frac{\sqrt{P_n/(2\pi\alpha)}}{r_n} e^{i\phi_n} e^{ikr_n}. \tag{7}$$

The *combined* solution, therefore, will be

$$\begin{aligned} y(\vec{x}, t) &= y_1(\vec{x}, t) + \ldots + y_N(\vec{x}, t) \\ &= \text{Re}\left\{A_1(\vec{x})e^{-i\omega t} + A_N(\vec{x})e^{-i\omega t}\right\}, \\ &= \text{Re}\left\{\left(A_1(\vec{x}) + \ldots + A_N(\vec{x})\right) e^{-i\omega}\right\}, \end{aligned}$$

which is just simple harmonic motion with a complex amplitude which is the sum of each of the complex amplitudes,

$$\begin{aligned} A(\vec{x}) &= A_1(\vec{x}) + \ldots + A_N(\vec{x}) \\ &= \frac{\sqrt{P_1/(2\pi\alpha)}}{r_1} e^{i\phi_1} e^{ikr_1} + \ldots + \frac{\sqrt{P_N/(2\pi\alpha)}}{r_N} e^{i\phi_N} e^{ikr_N}, \end{aligned}$$

where we have used Eq. 7.

We can simplify our work dramatically if we now express everything in terms of the intensities that we would see from each slit if it were the only slit (Eq. 6), $I_n \equiv P_n/(2\pi r_n^2)$. To do this, note that

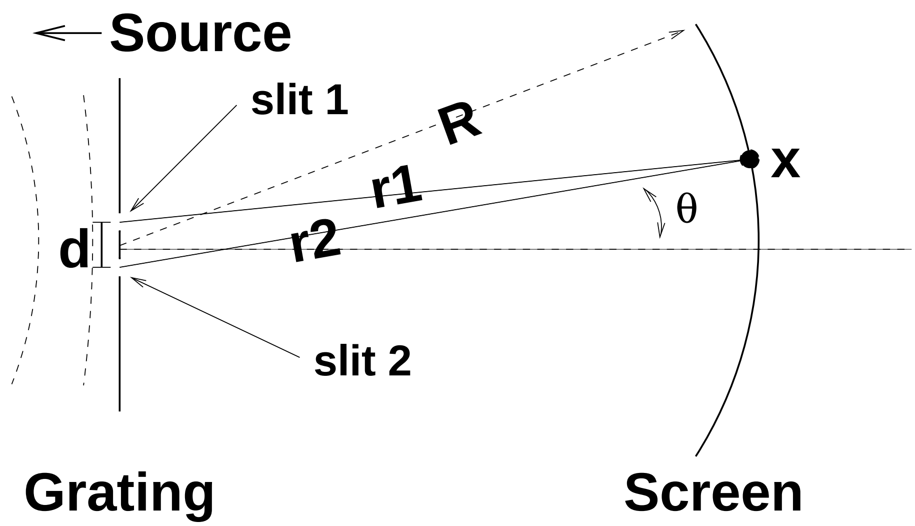

Figure 3: Double-slit interference experiment

$\sqrt{P_n/(2\pi\alpha)}/r_n = \sqrt{I_n/\alpha}$. Thus,

$$A(\vec{x}) \;=\; \sqrt{I_1/\alpha}e^{i\phi_1}e^{ikr_1} + \ldots + \sqrt{I_N/\alpha}e^{i\phi_N}e^{ikr_N}.$$

Finally, even the factors of α disappear when we compute the intensity from all of the slits combined,

$$\begin{aligned} I &= \alpha\,|A(\vec{x})|^2 \\ &= \alpha\left|\sqrt{I_1/\alpha}e^{i\phi_1}e^{ikr_1} + \ldots + \sqrt{I_N/\alpha}e^{i\phi_N}e^{ikr_N}\right|^2 \\ &= \left|\sqrt{\alpha}\left(\sqrt{I_1/\alpha}e^{i\phi_1}e^{ikr_1} + \ldots + \sqrt{I_N/\alpha}e^{i\phi_N}e^{ikr_N}\right)\right|^2 \\ &= \left|\sqrt{I_1}e^{i\phi_1}e^{ikr_1} + \ldots + \sqrt{I_N}e^{i\phi_N}e^{ikr_N}\right|^2, \end{aligned} \tag{8}$$

which is the simple result we gave at the start of this section. Note that most of this section has been managing algebra to get unknown constant factors to cancel out of the final expression, Eq. 8.

3.2 Two narrow slits

3.2.1 Analysis

We now consider the simplest case which exhibits interference, two equivalent, narrow slits (Figure 3). Of course, you can get the two-slit pattern by substituting $N = 2$ into the formula we shall eventually derive for N slits. But, deriving the two slit result as a special case first allows us to illustrate several useful results and techniques for complex arithmetic that you will have to master, in particular computing the amplitude of the complex number $\underline{c}$ with the formula $|\underline{c}|^2 = \underline{c}^*\underline{c}$. This section appears very mathematical because we show *all* of the steps. As you learn to do these steps, they will become simpler and simpler for you.

According to our main result, Eq. 8, the intensity at a point $\vec{x}$ on the screen for just two slits is

$$I \;=\; \left|\sqrt{I_1}e^{i\phi_1}e^{ikr_1} + \sqrt{I_2}e^{i\phi_2}e^{ikr_2}\right|^2, \tag{9}$$

where I_1 and I_2, ϕ_1 and ϕ_2, and r_1 and r_2, respectively, are (a) the intensities which each slit (if alone) would make at point $\vec{x}$, (b) the phases of the waves emerging from each slit, and (c) the distance from each slit to the point $\vec{x}$.

Now, because we assume the two slits to be equivalent, they must let the same intensity of waves through, and so the intensities which each separately would create on the screen are very nearly the same, so that $I_1 \approx I_2$. (There will be a slit difference in the distance from each slit to the point $\vec{x}$ on the screen, but so long as the screen is much further away than the separation of the slits, $R >> a$, this difference is negligible.) For simplicity, let us call this intensity which either slit alone creates on the screen I_0.

Next, because the waves coming from behind the diffraction grating arrive at both slits *in phase* (Figure 3), the phases of the waves coming out of the slits will be equal, $\phi_1 = \phi_2$. For simplicity, let us call this phase ϕ_0. As we shall see in a moment, the fact that these phases are equal means that they cancel from the final result. Note that this is only true if the waves coming from behind the grating arrive at the slits with equal phases!

Putting this all together we have,

$$\begin{aligned} I &= \left|\sqrt{I_0}e^{i\phi_0}e^{ikr_1} + \sqrt{I_0}e^{i\phi_0}e^{ikr_2}\right|^2 \\ &= \left|\sqrt{I_0}e^{i\phi_0}e^{ikr_1}\left(1+e^{ik(r_2-r_1)}\right)\right|^2. \\ &= \left|\sqrt{I_0}\right|^2\left|e^{i\phi_0}\right|^2\left|e^{ikr_1}\right|^2\left|1+e^{ik(r_2-r_1)}\right|^2. \\ &= I_0\left|1+e^{ik(r_2-r_1)}\right|^2, \end{aligned} \tag{10}$$

$$\tag{11}$$

where, again, we have used the facts that $|\underline{\mathrm{ab}}\ldots\underline{\mathrm{c}}|^2 = |\underline{\mathrm{a}}|^2|\underline{\mathrm{b}}|^2\ldots|\underline{\mathrm{c}}|^2$ and $|e^{i\phi}| = 1$ for any complex phase factor. Note that, as we had mentioned, in the end, the phases indeed cancel out. Also note that the entire result now depends only on the difference in lengths between the paths from the observation point to each slit, $\Delta r \equiv r_2 - r_1$.

The entire computation of I now boils down to finding $|1+e^{ik\Delta r}|^2$. A very useful way to find the square-magnitude of a complex number $\underline{\mathrm{c}}$ is to use the following fact

$$\begin{aligned} |\underline{\mathrm{c}}|^2 &= c_r^2 + c_i^2 \\ &= c_r^2 - (ic_i)^2 \\ &= (c_r - ic_i)(c_r + ic_i) \\ &\equiv \underline{\mathrm{c}}^*\underline{\mathrm{c}}, \end{aligned} \tag{12}$$

where, for any complex number $\underline{\mathrm{c}}$, we define the *complex conjugate* as $\underline{\mathrm{c}}^* \equiv c_r - ic_i$. Mathematically, taking the complex conjugate simply amounts to finding all occurrences of i and substituting $-i \to i$.

Using the procedure of Eq. 12, we now find

$$\begin{aligned} |1+e^{ik\Delta r}|^2 &= \left(1+e^{ik\Delta r}\right)^*\left(1+e^{ik\Delta r}\right) \\ &= \left(1+e^{-ik\Delta r}\right)\left(1+e^{ik\Delta r}\right) \\ &= 1 + e^{-ik\Delta r} + e^{ik\Delta r} + e^{-ik\Delta r}e^{+ik\Delta r} \\ &= 1 + (\cos(k\Delta r) - i\sin(k\Delta r)) + (\cos(k\Delta r) + i\sin(k\Delta r)) + 1 \\ &= 2 + 2\cos(k\Delta r) \\ &= 2\left(1+\cos(k\Delta r)\right) = 4\cos^2(k\Delta r/2), \end{aligned}$$

where on the last line we find a somewhat simpler form by using the trig identity $1+\cos(2\theta) = 2\cos^2(\theta)$. Combining this with Eq. 10, we have the final result for two-slits,

$$I = 4I_0\cos^2(k\Delta r/2). \tag{13}$$

All that remains is for us to learn to compute the *path difference* $\Delta r = r_2 - r_1$. Figure 4 shows that the path difference $r_2 - r_1$ corresponds very closely to the opposite side of a right triangle of angle θ and hypotenuse d. Accordingly, our final result is Eq. 13 with the definition

$$\Delta r = d\sin\theta. \tag{14}$$

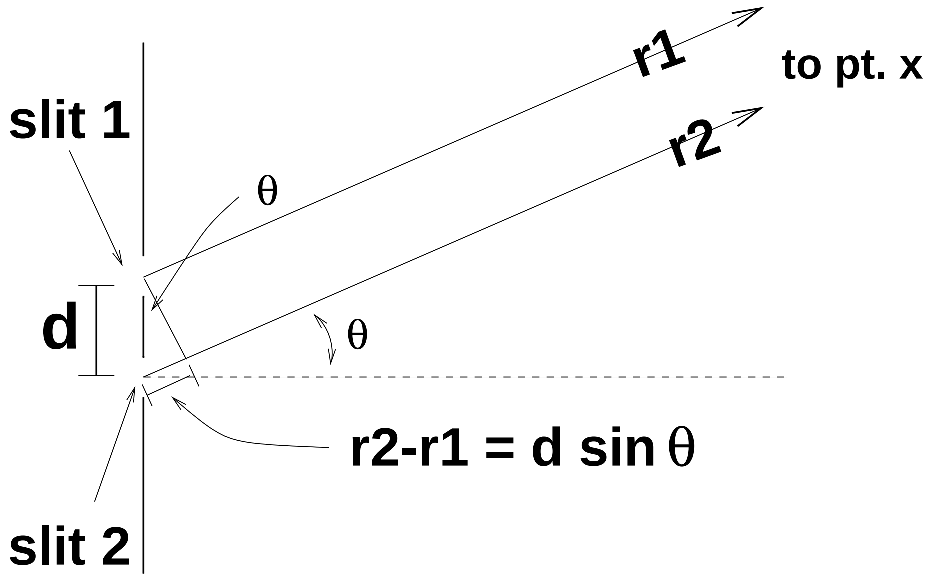

Figure 4: Double-slit interference experiment: enlarged view

3.2.2 Interpretation

Eqs. 13 & 14 give the intensity as a function of angle of observation. Figure 5 shows the resulting pattern. Three aspects of the result are noteworthy.

- The peak intensity, $4I_0$, is greater than the intensity we expect from simply summing the intensity of two slits. However, there are also places where the intensity drops to zero. In fact, *on average*, the intensity is exactly what we would expect from two slits, $2I_0$.

- The maxima in the pattern occur when $\cos(k\Delta r/2) = \pm 1$, which means $k\Delta r/2 = M\pi$, where M is some integer. This condition becomes then

$$\Delta r = M\frac{2\pi}{k} = M\lambda \quad \text{condition for two-slit maximum } (M \text{ an integer}), \tag{15}$$

  which means simply that the path difference is an even number of wavelengths so that the two waves arriving at $\vec{x}$ are in phase, adding together *constructively*. Under these conditions, the amplitude will then be exactly double, and because the intensity goes like the square of the amplitude, we expect 4 times the intensity, as Eq. 13 predicts.

- The minima (where there is zero intensity, actually) occur when $\cos(k\Delta r/2) = 0$, which means $k\Delta r/2$ is an odd half multiple of π, $k\Delta r/2 = (2M+1)\pi/2$, where M is any integer. This conditions then becomes

$$\Delta r = (M+1/2)\frac{2\pi}{k} = (M+1/2)\lambda \quad \text{condition for two-slit minimum } (M \text{ an integer}), \tag{16}$$

  which simply means that the path difference is off by exactly one-half of a wavelength, so that the two waves arrive *exactly* out of phase, and thus cancel completely, leaving zero amplitude and zero intensity.

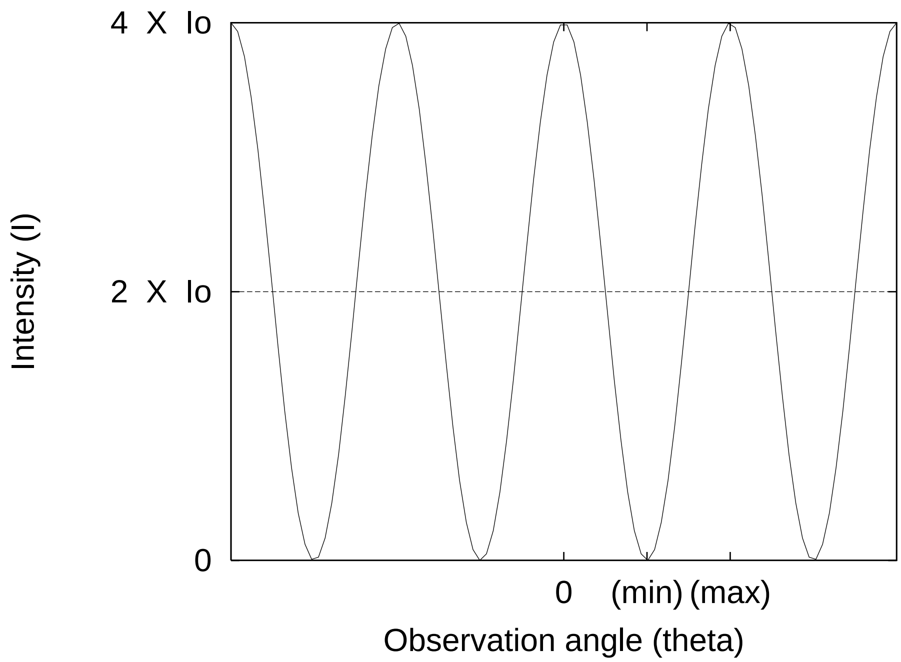

Figure 5: Intensity as a function of observation angle for a double-slit experiment

3.3 N narrow slits

3.3.1 Analysis

We now use all of the skills we have developed to derive what happens in the case of N identical, very narrow slits, all at equal separations d from one another (Figure 6). The only new technique in this section is that we use the formula for the sum of a geometric series.

Following our derivation for the two-slit case, we begin with

$$I = \left| \sqrt{I_0} e^{i\phi_0} e^{ikr_1} + \ldots + \sqrt{I_0} e^{i\phi_0} e^{ikr_N} \right|^2,$$

where we already have used the fact that all slits let through the same intensity (all I's are equal) and waves arrive at the slits and thus emerge from them all with the same phase (all ϕ's are the same).

Once again, we can factor out the constant phases, and the intensity we would expect from each slit individually, leaving only a sum of exponentials of path differences times k,

$$\begin{aligned} I &= \left| \sqrt{I_0} e^{i\phi_0} e^{ikr_1} \left(1 + \ldots + e^{ik(r_N - r_1)} \right) \right|^2, \\ &= \left| \sqrt{I_0} \right|^2 \left| e^{i\phi_0} \right|^2 \left| e^{ikr_1} \right|^2 \left| 1 + \ldots + e^{ik(r_N - r_1)} \right|^2, \\ &= I_0 \left| 1 + e^{ik(r_2 - r_1)} + e^{ik(r_3 - r_1)} \ldots + e^{ik(r_N - r_1)} \right|^2. \end{aligned}$$

The next key step is to realize that the path differences are just multiples of each other. If we define, again, $\Delta r = r_2 - r_1 = d \sin\theta$, then $r_3 - r_1 = 2\Delta r$, $r_4 - r_1 = 3\Delta r$, …, $r_N - r_1 = (N-1)\Delta r$. (See Figure 6.) Thus, we must compute the sum

$$1 + e^{ik\Delta r} + e^{2ik\Delta r} + e^{3ik\Delta r} + \ldots e^{(N-1)ik\Delta r},$$

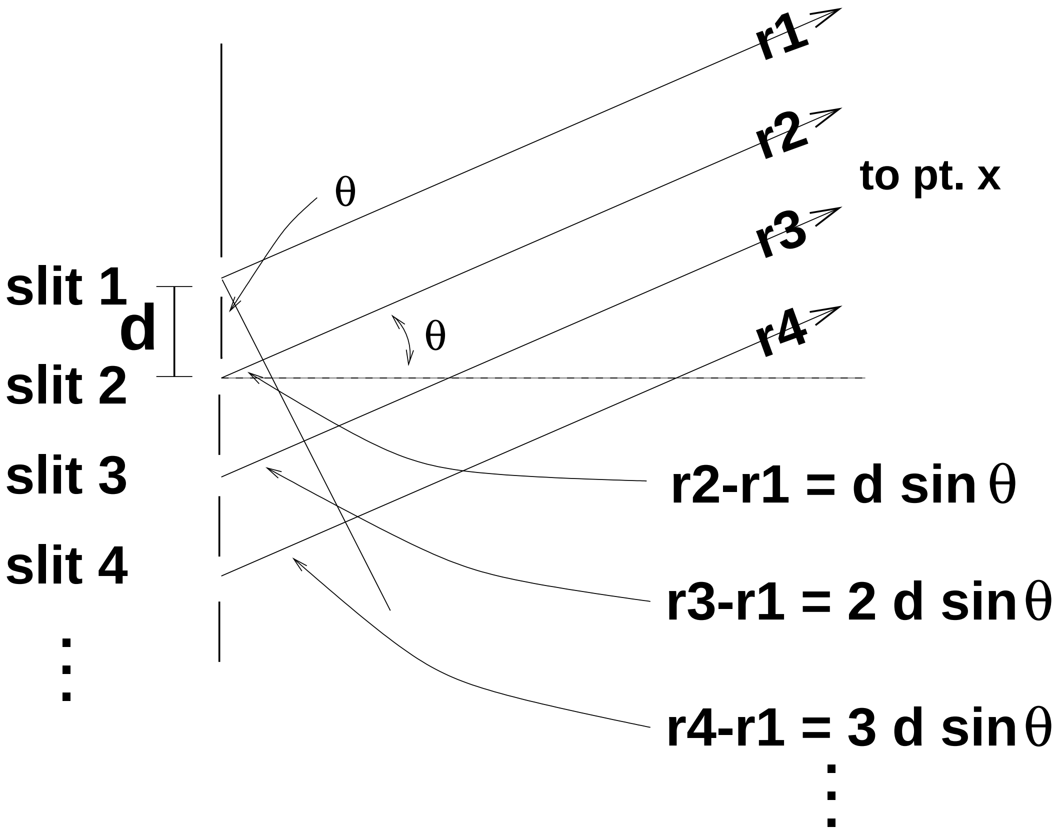

Figure 6: Interference experiment with N finite slits

but this is just a geometric series $G = 1 + p + p^2 + \ldots + p^{N-1}$ with geometric ratio $p = e^{ik\Delta r}$. The formula you know for real numbers, $G = 1/(1-r)$ still works, and so the sum is

$$1 + e^{ik\Delta r} + e^{2ik\Delta r} + e^{3ik\Delta r} + \ldots e^{(N-1)ik\Delta r} = \frac{(e^{ik\Delta r})^N - 1}{e^{ik\Delta r} - 1} = \frac{e^{iNk\Delta r} - 1}{e^{ik\Delta r} - 1}.$$

To compute the magnitude of this, we use the fact that the magnitude of the ratio is the ratio of the magnitudes. Then, we note that both numerator and denominator have the same form, $|1 - e^{i\Phi}|^2$, where for the former case $\Phi = Nk\Delta r$ and for the latter $\Phi = k\Delta r$. For both cases, we have

$$\begin{aligned}
|1 - e^{i\Phi}|^2 &= \left(1 - e^{i\Phi}\right)^* \left(1 - e^{i\Phi}\right) \\
&= \left(1 - e^{-i\Phi}\right)\left(1 - e^{i\Phi}\right) \\
&= 1 - e^{-i\Phi} - e^{i\Phi} + 1 \\
&= 2 - 2\cos(\Phi) \\
&= 2\left(1 - \cos(\Phi)\right) = 4\sin^2(\Phi/2),
\end{aligned}$$

where, again, at the last step we use a convenient trig identity. Using this, our final result is

$$\begin{aligned}
I &= I_0|1 + e^{ik(r_2 - r_1)} + e^{ik(r_3 - r_1)} \ldots + e^{ik(r_N - r_1)}|^2. \\
&= I_0 \left| \frac{e^{iNk\Delta r} - 1}{e^{ik\Delta r} - 1} \right| \\
&= I_0 \frac{\left|e^{iNk\Delta r} - 1\right|^2}{\left|e^{ik\Delta r} - 1\right|^2} \\
&= I_0 \frac{4\sin^2(Nk\Delta r/2)}{4\sin^2(k\Delta r/2)},
\end{aligned}$$

which simplifies to our final result,

$$I = I_0 \frac{\sin^2(Nk\Delta r/2)}{\sin^2(k\Delta r/2)}, \tag{17}$$

where, as before, $\Delta r = d\sin\theta$.

3.3.2 Interpretation

Again, several features of our result, Eq. 17, are noteworthy. Figure 7 shows this result for a case with $N = 6$ slits.

- Case of two slits: If we substitute $N = 2$ into Eq. 17 and use the double-angle sine formula, we have for the numerator $\sin(2k\Delta r/2) = 2\sin(k\Delta r/2)\cos(k\Delta r/2)$, so that the $\sin^2$ term in the denominator cancels, leaving $4I_0\cos^2(k\Delta r/2)$, precisely the two-slit result, Eq. 13.
- Principle maxima: we find extremely large maxima whenever the denominator in Eq. 13 goes to zero. This occurs whenever $\sin(k\Delta r/2) = 0$, equivalently, $k\Delta r/2 = M\pi$, where M is some integer. This means
$$\Delta r = M(2\pi/k) = M\lambda \ \text{ condition for } N\text{-slit principle maximum } (M \text{ an integer}),$$
so that the path difference between *all* of the slits is always an exact number of wavelengths, leading to complete constructive interference. To see how large the maxima can be, we take the limit $\Delta r \longrightarrow 0$. From the small-angle formula, we then have
$$I \approx I_0 \frac{(Nk\Delta r/2)^2}{(k\Delta r/2)^2} = I_0 N^2 \ \text{ Intensity of } N\text{-slit principle maximum.}$$
The physical reason for the factor of N^2 is just that the amplitudes of all N waves add, so the amplitude is N times as large, but the intensity, of course, goes like the amplitude *squared*.

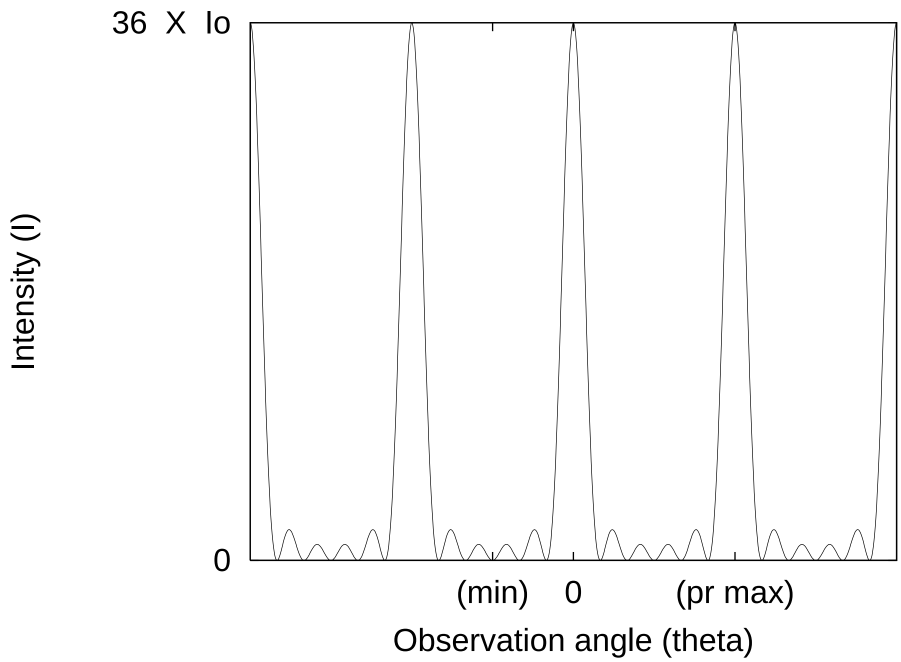

Figure 7: Intensity as a function of observation angle for an N-slit experiment for the case N=6. In this case, principle maxima appear in the location of every sixth minima, so that there are five minima between principle maxima, and thus four lesser maxima between principle maxima.

- Minima: Minima occur in the pattern whenever the numerator is zero (and the denominator is not). Zeros in the numerator occur whenever $\sin(Nk\Delta r/2) = 0$, equivalently, $Nk\Delta r/2 = M\pi$, where M is any integer, so that

  $\Delta r = (M/N)(2\pi/k) = (M/N)\lambda$ condition for N-slit minimum (M a non-zero integer, M/N *not* an integer).

  Note that the spacing between minima is thus exactly $1/N$ times the spacing between principle maxima. Often, you can use this to determine the number of slits from the appearance of the diffraction pattern.

- Lesser maxima: Of course, there are other maxima which must appear approximately halfway between neighboring minima, the M^{th} of these happens where

  $$\Delta r = \frac{M+1/2}{N}(2\pi/k) = \frac{M+1/2}{N}\lambda \ \text{ condition for } M^{\text{th}}\ N\text{-slit lesser maxima } (M \text{ an integer}).$$

4 Finite slit: Diffraction

Figure 8 shows how we can compute the intensity pattern for a slit of finite width a (no longer infinitesimal compared to the wavelength) by using our result for N slits. The idea is simple. If we made a diffraction grating with N small slits spaced very closely apart at distance $d = a/N$, and kept adding more little slits, eventually we end up removing all of the grating, leaving just a single slit of width $N \cdot d = N(a/N) = a$. The special term *diffraction* is generally used to describe interference patterns formed by finite openings. The underlying physical principle, however, remains interference.

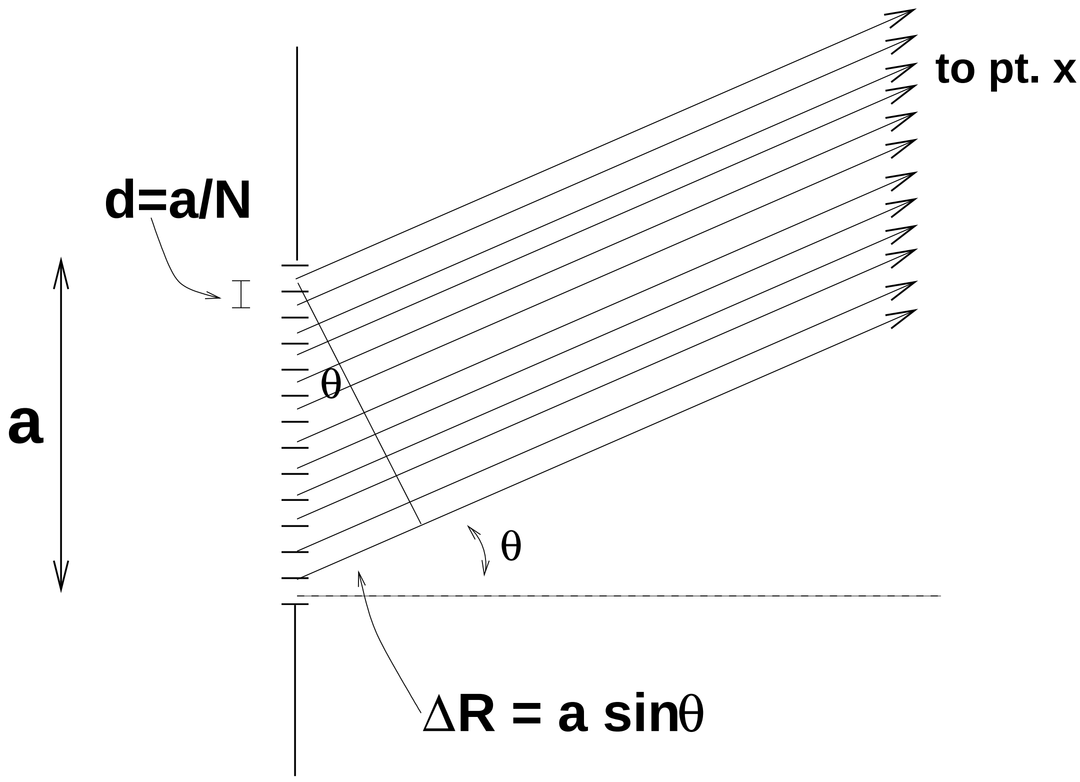

Figure 8: Analysis of diffraction experiment with slit of finite width

4.1 Analysis

According to the physical reasoning of the preceding paragraph, we obtain the diffraction pattern of a single slit by substituting $d = a/N$ into Eq. 17 (so that we always have the correct total width a) and then taking the limit $N \longrightarrow \infty$ (so that we have in the end a single, continuous slit). Recalling that $\Delta r = d\sin\theta$, the factor in numerator then becomes $Nk\Delta r/2 = Nk(a/N\sin\theta)/2 = ka\sin\theta/2$. As a similar term appears in the denominator, it is convenient to define the path difference between the top and bottom of the slit (Figure 8)

$$\Delta R \equiv a\sin\theta. \tag{18}$$

With this, the phase factor in the numerator becomes $\sin^2(k\Delta R/2)$. Finally, noting that the argument of the sine function in the denominator is the same but for a missing factor of N, the denominator is just $\sin^2(k\Delta R/(2N))$.

This leaves for us to take the $N \longrightarrow \infty$ limit of

$$I = I_0 \frac{\sin^2(k\Delta R/2)}{\sin^2(k\Delta R/(2N))}.$$

As $N \longrightarrow \infty$, the argument of the sine function in the denominator becomes quite small, so that we may use the small angle formula, $\sin(k\Delta R/(2N)) \approx k\Delta R/(2N)$. With this, we have

$$I = I_0 N^2 \frac{\sin^2(k\Delta R/2)}{(k\Delta R/2)^2}.$$

From the discussion in Section 3.3.2, we know the intensity at the central maximum to be $I_0 N^2$. Defining this intensity as $I_{max} \equiv I_0 N^2$, we have our final result for the single slit,

$$I = I_{max} \frac{\sin^2(k\Delta R/2)}{(k\Delta R/2)^2}, \tag{19}$$

where, recall, $\Delta R \equiv a\sin\theta$.

4.2 Interpretation

Figure 9 shows an example of a diffraction pattern from a single slit. Again, several features are noteworthy.

- A single, central "principle maximum": From the discussion for the N-slit pattern, we only see principle maxima after every $N-1$ minima, where N is the number of slits. Because we now have an effectively infinite number of slits, we only see the central principle maximum at $\theta = 0$.

- Minima: Minima occur whenever the numerator is zero (and the denominator is not). This happens whenever $\sin(k\Delta R/2) = 0$, equivalently

  $$k\Delta R/2 = M\pi \quad \text{condition for finite-slit minimum } (M \text{ any non-zero integer} - \text{to avoid central max})$$

  This corresponds to $\Delta R = M(2\pi/k) = M\lambda$, a path difference between one end of the slit and the other of any non-zero integer number of wavelengths. If the path difference is zero, all the contributions interference constructively, and you get a large maximum. If the path difference is not zero, but an even number of wavelengths, then we see that you always find exactly canceling negative and positive contributions from difference regions of the slit, leading ultimately to a minimum.

- Lesser Maxima: Again, there will be lesser maxima situated approximately halfway in between the minima.

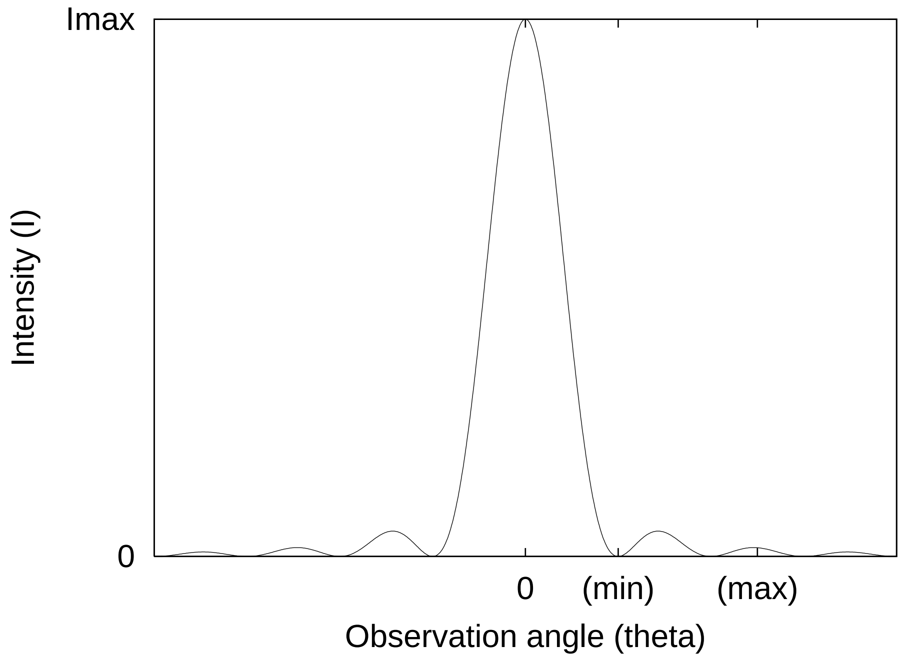

Figure 9: Intensity as a function of observation angle for a finite-slit experiment

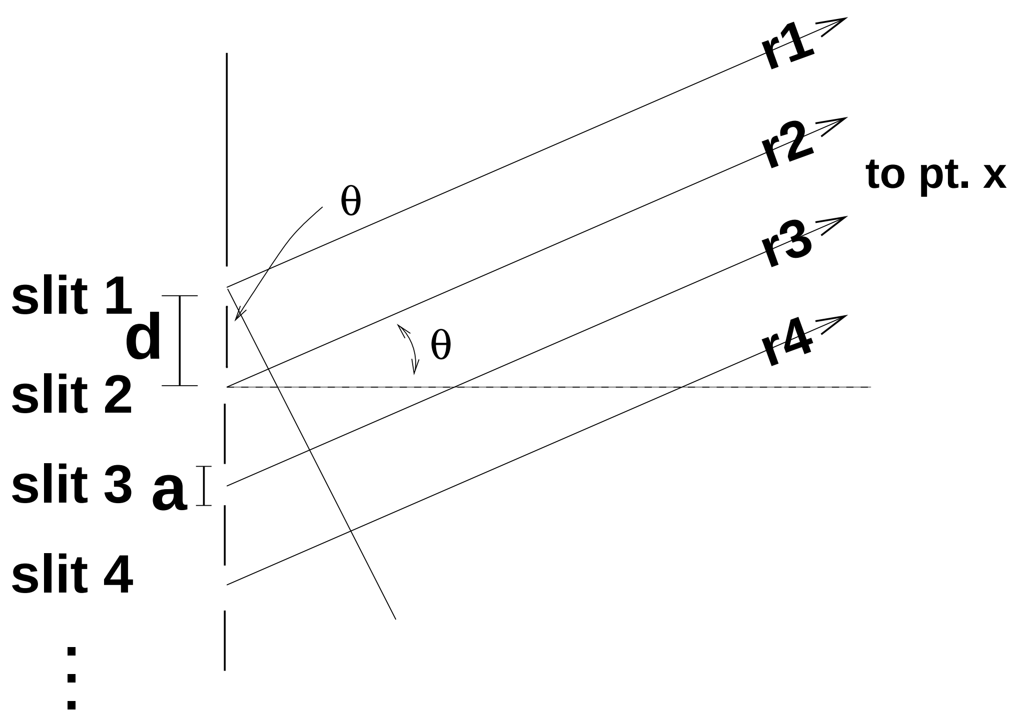

Figure 10: N finite-slit interference experiment

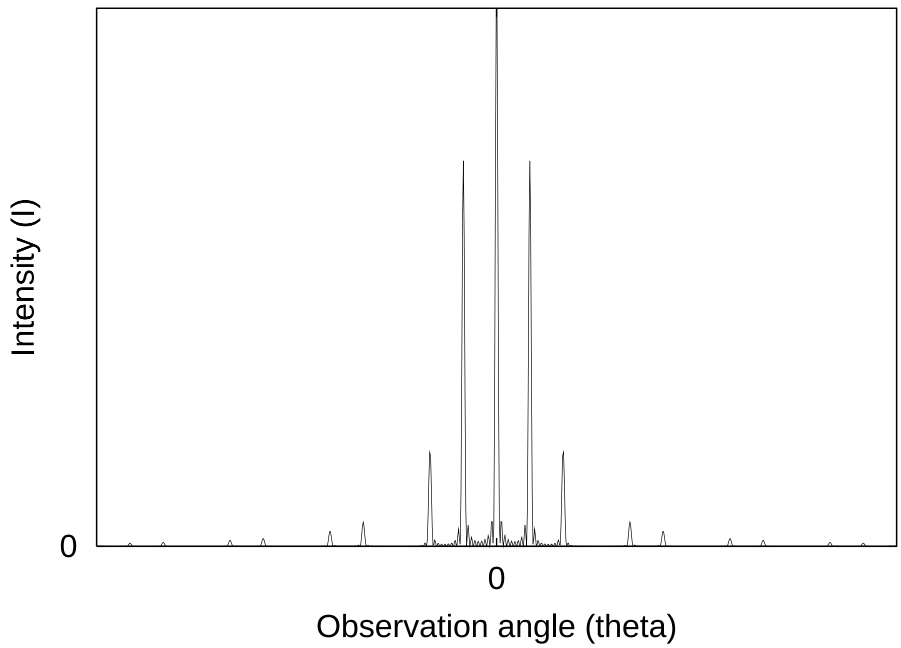

Figure 11: Intensity as a function of observation angle for an experiment with N slits of finite width, where N is large ($N = 10$) and the slits are spaced four times further apart than they are wide, $d/a = 3$. Note how this pattern has every third principle maxima (except the central maximum) wiped away.

5 N finite slits: combining diffraction and interference

5.1 Analysis

Finally, we consider the pattern created by N identical slits at spacing d, but each with finite width a (Figure 10). As you might expect, this case becomes a simple combination of the result for the finite slit with the result for N slits.

The reasoning for this is as follows. If you view any one of the N finite slits, say slit n, *by itself* from different angles, it will appear to change its intensity according to the single-slit formula $I_n(\theta) = I_{max} \sin^2(k\Delta r/2)/(k\Delta r/2)^2$, where $\Delta R = a \sin\theta$. If we now open all N slits at once and view the result at angle θ, what we see is the result of the interference of N slits, each with this same intensity. Thus, if we set I_0 from Eq. 17 equal to $I_{max} \sin^2(k\Delta r/2)/(k\Delta r/2)^2$, we will find the result for N finite slits. This result is just the *product* of our two previous results:

$$I = I_{max} \frac{\sin^2(k\Delta R/2)}{(k\Delta R/2)^2} \frac{\sin^2(Nk\Delta r/2)}{\sin^2(k\Delta r/2)}, \tag{20}$$

where $\Delta R = a \sin\theta$, $\Delta r = d \sin\theta$, and I_{max} is the central maximum intensity from any one of the slits alone.

5.2 Interpretation

Figure 11 shows the result for a case where the spacing between slits d is exactly three times the width of the slits. Among many interesting features, two are noteworthy in this result.

- Spacing between maxima: Because Δr is larger (d, the spacing between slits, must be larger than a, the width of the slits, otherwise the slits overlap!), the most rapid oscillations come from the N-slit interference pattern, which then sets the basic pattern. Note that if N is large (N is only ten in the figure), then you see mostly the principle maxima. The little fuzz in the figure hardly visible near zero *are* the lesser maxima and would become nearly invisible for any large value of N. This is a general feature of the N slit pattern and not particular to the case of finite slits!

- Missing Principle Maxima: The main impact from the finite width of the slits is that it changes the amplitude of the principle maxima. Rather than having equal heights, the height of each maximum is multiplied by the finite-slit factor in Eq. 20. We expect the N-slit pattern (Section 3.3.2) to have principle maxima wherever $d \sin\theta = \Delta r = M\lambda$, where M is an integer, so that $\sin\theta = M\frac{\lambda}{d}$. However, these maxima can sometimes be wiped away by the finite-slit diffraction factor in Eq. 20. In particular, the zeros of the finite-slit diffraction pattern occur where $a \sin\theta = \Delta R = M'\lambda$, or $\sin\theta = M'\frac{\lambda}{a}$, where M' is another integer. Whenever both of these conditions are met at once, we find that the M^{th} expected maximum appears to be missing.

  Equating $\sin\theta$ in both cases, we find that this happens when

  $$M\frac{a}{d} = M'$$

  is an integer. In our example, $a/d = 1/3$, and so we expect every third principle maximum M (except the one in the center, $M = 0$) to be missing, exactly as in Figure 11!